AD-A072 432 MISSOURI UNIV-COLUMBIA DEPT OF STATISTICS F/G 12/1

PRODUCT INEQUALITIES INVOLVING THE MULTIVARIATE NORMAL DISTRIBU--ETC(U)

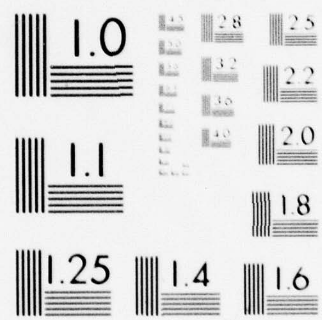
JUL 79 R L DYKSTRA N00014-78-C-0655

UNCLASSIFIED TR-85 NL

| OF |

AD
A072432

END
DATE
FILMED
9-79
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

LEVEL II

University of Missouri-Columbia

Product Inequalities Involving the Multivariate Normal Distribution

BY

R. L. Dykstra

Technical Report No. 85
Department of Statistics
July 1979

Mathematical Sciences

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 85	2. GOVT ACCESSION NO. (14) TR-85	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) (6) Product Inequalities Involving the Multivariate Normal Distribution.		5. TYPE OF REPORT & PERIOD COVERED (9) Technical Report.
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) Richard L. Dykstra		8. CONTRACT OR GRANT NUMBER(s) (15) N00014-78-C-0655
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics University of Missouri Columbia, MO 65211		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Office of Naval Research Department of the Navy Arlington, VA		12. REPORT DATE (11) July 1979
		13. NUMBER OF PAGES 20 pages
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (12) 23p.		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Product inequality, multivariate normal, Wishart matrix, simultaneous confidence intervals in multivariate regression.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Suppose $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k')$ possesses a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and positive semidefinite covariance matrix Σ. If $C_i \subset R^{p_i}$ denote convex regions symmetric about the origin, then conditions are given such that

DD FORM 1473, 1 JAN 73 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

Unclassified 402 660

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

79 08 7 010

20.

$$P(\underset{\sim}{Y}_i \in C_i,\ i = 1,\ldots,k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in C_i)$$

and/or

$$P(Y_i \in \overline{C}_i,\ i = 1,\ldots,k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in \overline{C}_i)$$

obtain. These conditions imply that chi-squared random variables defined from a multivariate normal distribution are always positively dependent and non negatively correlated. Other applications involve conservative simultaneous confidence regions in a multivariate regression setting.

Product Inequalities Involving the
Multivariate Normal Distribution

by

R. L. Dykstra

Technical Report No. 85

July 1979

Prepared under contract N00014-78-C-0655
for the Office of Naval Research

Department of Statistics
University of Missouri
Columbia, Missouri 65211

Product Inequalities Involving the

Multivariate Normal Distribution

Abstract

Suppose $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k')$ possesses a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and positive semidefinite covariance matrix Σ. If $C_i \subset R^{p_i}$ denote convex regions symmetric about the origin, then conditions are given such that

$$P(\underset{\sim}{Y}_i \in C_i,\ i = 1, \ldots, k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in C_i)$$

and/or

$$P(Y_i \in \overline{C}_i,\ i = 1, \ldots, k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in \overline{C}_i)$$

obtain. These conditions imply that chi-squared random variables defined from a multivariate normal distribution are always positively dependent and non negatively correlated. Other applications involve conservative simultaneous confidence regions in a multivariate regression setting.

AMS Classification numbers: Primary 62H99; Secondary 62E10.

Key words and phrases: product inequality, multivariate normal, Wishart matrix, simultaneous confidence intervals in multivariate regression.

1. INTRODUCTION.

The following question has received the attention of several investigators. Suppose $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1', \underset{\sim}{Y}_2', \ldots, \underset{\sim}{Y}_k')$ possesses a multivariate normal distribution with mean vector $\underset{\sim}{0}$ and positive semidefinite (p.s.d.) covariance matrix Σ, i.e. $\underset{\sim}{Y}$ is $n(\underset{\sim}{0}, \Sigma)$. The dimension of $\underset{\sim}{Y}_i$ is p_i so that the dimension of $\underset{\sim}{Y}$ must be $\sum_1^k p_i$. We let C_i denote a convex region in R^{p_i} (and $\overline{C}_i$ its complement) which is symmetric about the origin ($\underset{\sim}{x} \in C_i$ implies $-\underset{\sim}{x} \in C_i$). Under what conditions will at least one of the two inequalities

$$(1.1) \qquad P(\underset{\sim}{Y}_i \in C_i,\ i = 1, \ldots, k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in C_i)$$

and

$$(1.2) \qquad P(\underset{\sim}{Y}_i \in \overline{C}_i,\ i = 1, \ldots, k) \geq \prod_{i=1}^{k} P(\underset{\sim}{Y}_i \in \overline{C}_i)$$

hold? A primary application of such inequalities has been obtaining conservative simultaneous confidence bounds which are easily computable although the inequality is useful in other areas as well.

Dunn (1958) showed the inequality (1.1) would hold if $p_i = 1$ for all i and the C_i were equal, providing $k = 2$ or $k = 3$. (For $k = 2$ it is easily shown that (1.1) and (1.2) are equivalent.) Moreover she proved her conjecture for arbitrary k if the correlation matrix corresponding to Σ was of the form $\rho_{ij} = b_i b_j$, $i,j = 1,\ldots,k$, $i \neq j$ and $0 < b_i < 1$, $i = 1,\ldots,k$.

Sidak (1967) was able to extend Dunn's results to the case of an arbitrary covariance matrix and different C_i's (although still one dimensional).

Scott (1967) purported to prove the same result as Sidak as well as inequality (1.2) with the same conditions holding. However a subtle conditioning error occurs in Scott's arguments and invalidates his proofs. Moreover, Sidak (1971) has constructed a counterexample to Scott's second inequality so that (1.2) does not hold under these conditions.

In light of Sidak's 1967 results it is tempting to conjecture an analogue of a one sided result due to Slepian (1962), namely that decreasing the absolute value of the correlations should decrease $P(c_i \le Y_i \le c_i,\ i = 1, \ldots, k)$. However a counter example of Sidak (1968) showed that this is not true in general. He did show however that the conjecture held if the absolute values of the correlations decreased in a particular manner.

Khatri (1967) was able to prove (1.1) for higher dimensions (although the p_i must be equal) when the C_i's are ellipsoids of the form $\{\underset{\sim}{x} \cdot \underset{\sim}{x}'A\underset{\sim}{x} \le c_i\}$, the covariance structure of $\underset{\sim}{Y}$ is the Kronecker product $\Psi^{k\times k} \otimes I^{p\times p}$, and A and Ψ are positive semidefinite.
By adding a restriction similar to Dunn's on the matrix Ψ , Khatri was able to prove (1.2) in the same setting. He also considered some situations where the ellipsoids are random. Later Khatri (1970) had apparently extended his

1967 results to the point that (1.1) and (1.2) hold without any restrictions on Σ or the C_i's (other than those initially assumed). However Sidak (1975) isolated an unobtrusive error in both Scott's 1967 paper and Khatri's 1970 paper. Moreover, Sidak's 1971 counterexample has shown that the general inequality (1.2) without restrictions is incorrect. Das Gupta et. al. (1972) considered inequalities of the type (1.1) and (1.2) in the more general setting of elliptically contoured distributions. They were able to establish inequality (1.1) with out restrictions on Σ providing every p_i exceptone equals one.

Tong (1970) and Sidak (1973) have established some inequalities related to (1.1) and (1.2) when all the correlations determined by Σ are identical.

Whether inequality (1.1) holds in general without additional assumptions has been an open question discussed in several of the above references. It is the purpose of this paper to give some new conditions under which inequalities (1.1) or (1.2) hold and discuss briefly a few applications which result.

2. THEOREMS AND PROOFS.

We will have need of the following lemmas in proving our theorems.

LEMMA 1. If $\underset{\sim}{Z}' = (Z_1, \ldots, Z_p)$ is a random vector with density $f(\underset{\sim}{z})$ symmetric about $\underset{\sim}{0}$ $(f(\underset{\sim}{z}) = f(-\underset{\sim}{z})$ for all

$\underset{\sim}{z}$) such that $\{\underset{\sim}{z};\ f(\underset{\sim}{z}) \le c\}$ is convex for all non-negative c, $\underset{\sim}{a}$ is an arbitrary fixed vector in R^p, and C is a convex set symmetric about the origin, then $P(\underset{\sim}{Z} \in C + \lambda\underset{\sim}{a})$ is a nonincreasing function of $\lambda (0 \le \lambda < \infty)$.

This lemma is due to T. A. Anderson (1955) and is well known.

LEMMA 2. Suppose the square, symmetric matrix A is partitioned as $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ where A_{11} and A_{22} (nonsingular) are square. Then A is positive definite (p.d.) iff both $A_{11} - A_{12}A_{22}^{-1}A_{21}$ and A_{22} are positive definite.

PROOF. It is well known that if E is a p.d. matrix and D is nonsingular (same dimension), then DED' is p.d. If we define the matrix

$$B = \begin{pmatrix} I & -A_{12}A_{22}^{-1} \\ 0 & I \end{pmatrix},$$

then B is clearly nonsingular. However then

$$BAB' = \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & 0 \\ 0 & A_{22} \end{pmatrix},$$

which implies the desired result.

LEMMA 3. If $h_1(x)$, $h_2(x)$ are nonnegative functions either both nonincreasing or both nondecreasing and X is a random variable, then

$$E\, h_1(X)\, h_2(X) \geq E\, h_1(X) \cdot E\, h_2(X) .$$

This result is well known and appears in various places, one place being Kimball (1951).

LEMMA 4. If $\underset{\sim}{x} = (x_1, \ldots, x_p)'$ is a p-variate vector and $\phi(\underset{\sim}{x})$ is a real-valued function invariant under orthogonal transformation $(\phi(A\underset{\sim}{x}) = \phi(\underset{\sim}{x})$ for every orthogonal matrix A), then $\phi(\underset{\sim}{x})$ is a function of $\underset{\sim}{x}$ only through Σx_i^2.

PROOF. For a given vector $\underset{\sim}{x} = (x_1, \ldots, x_p)'$, let $0(\underset{\sim}{x})$ denote an orthogonal matrix such that $0(\underset{\sim}{x})\, \underset{\sim}{x} = ((\Sigma x_i^2)^{\frac{1}{2}}, 0, \ldots, 0)'$. (WLOG we may assume $0(a\underset{\sim}{x}) = 0(\underset{\sim}{x})$ for any positive constant a.) The assumption then guarantees that $\phi(\underset{\sim}{x}) = \phi(0(\underset{\sim}{x})\underset{\sim}{x})$ is a function of only Σx_i^2.

Theorem 1 is, in one sense, a generalization of a theorem by Das Gupta et. al. (1973).

<u>THEOREM 1</u>. Suppose $\begin{pmatrix} \underset{\sim}{Y}_1 \\ \underset{\sim}{Y}_2 \end{pmatrix}$ is $n(\underset{\sim}{0}, \Sigma)$ with covariance matrix Σ expressable as $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & I_{22} \end{pmatrix}$ and that C is an arbitrary convex set symmetric about $\underset{\sim}{0}$. Then if Λ is idempotent $(\Lambda^2 = \Lambda)$,

$$P(\underset{\sim}{Y}_1 \varepsilon C,\ \underset{\sim}{Y}_2' \Lambda \underset{\sim}{Y}_2 \leq c) \geq P(\underset{\sim}{Y}_1 \varepsilon C) P(Y_2' \Lambda Y_2 \leq c) \tag{2.1}$$

for all $c \geq 0$.

PROOF. By making linear transformations on $\underset{\sim}{Y}_1$ and $\underset{\sim}{Y}_2$ which may reduce the number of random variables, we may assume WLOG that both Σ_{11} and I_{22} are of full rank and that the left side of (2.1) is

$$P(\underset{\sim}{Y}_1 \in C \ , \ \underset{\sim}{Y}_2' \underset{\sim}{Y}_2 \leq c) \ .$$

We may also assume WLOG that Σ is full rank, since if it is not, we may define a sequence of positive definite matrices of the same form which converge elementwise to Σ and then use a limiting argument to obtain the desired conclusion.

By a continuity argument, we may replace I_{22} by $(1-\varepsilon)I_{22}$ and still have Σ be positive definite if ε is sufficiently small. Thus the augmented matrix,

$$\Sigma^* = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & \Sigma_{12} \\ \Sigma_{21} & I_{22} & (1-\varepsilon)I_{22} \\ \Sigma_{21} & (1-\varepsilon)I_{22} & (1-\varepsilon)I_{22} \end{pmatrix} , \tag{2.2}$$

is positive definite for ε sufficiently small.

To see this, partition off the last row and column of submatrices and note from Lemma 2 that Σ^* is p.d. iff

$$\begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{21}/(1-\varepsilon) & 0 \\ 0 & \varepsilon I_{22} \end{pmatrix} \tag{2.3}$$

and $(1-\varepsilon)I_{22}$ are p.d.

Suppose now that $(\underset{\sim}{Y}_1', \underset{\sim}{Y}_2', \underset{\sim}{Z}')'$ is $n(\underset{\sim}{0}, \Sigma^*)$. Then we may write

$$P(\underset{\sim}{Y}_1 \varepsilon C,\ \underset{\sim}{Y}_2'\underset{\sim}{Y}_2 \le c) = \underset{\underset{\sim}{Z}}{E} P(\underset{\sim}{Y}_1 \varepsilon C,\ \underset{\sim}{Y}_2'\underset{\sim}{Y}_2 \le c \mid \underset{\sim}{Z})$$

$$(2.4) \qquad = \underset{\underset{\sim}{Z}}{E}\{P(\underset{\sim}{Y}_1 \varepsilon C \mid \underset{\sim}{Y}_1 \text{ is } n(\Sigma_{12}\underset{\sim}{Z}/(1-\varepsilon),\ \Sigma_{11} - \Sigma_{12}\Sigma_{21}/(1-\varepsilon)) \cdot$$

$$P(\underset{\sim}{Y}_2'\underset{\sim}{Y}_2 \le c \mid \underset{\sim}{Y}_2 \text{ is } n(\underset{\sim}{Z},\ \varepsilon I_{22})\}$$

where $\underset{\sim}{Z}$ is $n(\underset{\sim}{0},(1-\varepsilon)I_{22})$. If we let $|\underset{\sim}{Z}| = (\sum_i Z_i^2)^{\frac{1}{2}}$ denote the distance from $\underset{\sim}{Z}$ to the origin, we may express $\underset{\sim}{Z} = Q(|\underset{\sim}{Z}|, 0, \ldots, 0)'$ where the random orthogonal matrix Q and $|\underset{\sim}{Z}|$ are independent. Moreover, since the distribution of Q is invariant if multiplied on the left by a fixed orthogonal matrix, Q must posess the Haar invariate distribution discussed in Anderson (1958). Since the conditional distribution of $\underset{\sim}{Y}_2'\underset{\sim}{Y}_2$ is (except for a constant) a noncentral chi squared distribution with noncentrality parameter $|\underset{\sim}{Z}|^2$, and hence free of Q, we may express the quantity in (2.4) as

$$(2.5) \qquad \underset{|Z|}{E}\,[\underset{Q}{E}\ P(\underset{\sim}{Y}_1 \varepsilon C \mid Q, |\underset{\sim}{Z}|)] \cdot P(\underset{\sim}{Y}_2'\underset{\sim}{Y}_2 \le c \mid |\underset{\sim}{Z}|)] \ .$$

By Lemma 1, the second factor of the integrand is a decreasing function of $|\underset{\sim}{Z}|$. Similarly, for a fixed value of Q, the integrand of $\underset{Q}{E}$ is nonincreasing in $|\underset{\sim}{Z}|$, and hence $\underset{Q}{E}\, P(\underset{\sim}{Y}_1 \varepsilon C | Q, |\underset{\sim}{Z}|)$ is nonincreasing in $|\underset{\sim}{Z}|$. Then by applying Lemma 3, the expression in (2.5) must be

$$\geq \underset{|\underset{\sim}{Z}|,Q}{E}\, P(\underset{\sim}{Y}_1 \varepsilon C | Q, |\underset{\sim}{Z}|) \cdot \underset{|\underset{\sim}{Z}|}{E}\, P(\underset{\sim}{Y}_2' \underset{\sim}{Y}_2 \leq c \,|\, |\underset{\sim}{Z}|)$$

$$= P(\underset{\sim}{Y}_1 \varepsilon C) \cdot P(\underset{\sim}{Y}_2' \underset{\sim}{Y}_2 \leq c)$$

which was the desired result.

By replacing the words "nonincreasing in $|\underset{\sim}{Z}|$" by "nondecreasing in $|\underset{\sim}{Z}|$" the following corollary is immediate.

<u>COROLLARY 1</u>. Under the same assumptions as in Theorem 1,

$$P(\underset{\sim}{Y}_1 \varepsilon \overline{C},\ \underset{\sim}{Y}_2' A \underset{\sim}{Y}_2 \geq c) \geq P(\underset{\sim}{Y}_1 \varepsilon \overline{C}) P(\underset{\sim}{Y}_2' A \underset{\sim}{Y}_2 \geq c) .$$

If we assume that C is of the form

$$C = \bigcap_i \{\underset{\sim}{Y},\ \underset{\sim}{Y}_i' \Lambda_i \underset{\sim}{Y}_i \leq c_i\}$$

where the Λ_i are idempotent matrices, then repeated application of Theorem 1 gives the following corollary.

COROLLARY 2. If $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k')$ is $n(\underset{\sim}{0},\Sigma)$ where the corresponding partition of Σ is

$$\Sigma = \begin{pmatrix} I_{11} & \Sigma_{12} & \cdots & \Sigma_{1k} \\ \Sigma_{21} & I_{22} & & \vdots \\ \vdots & & \ddots & \vdots \\ \Sigma_{k1} & & \cdots & I_{kk} \end{pmatrix}$$

then

$$(2.6) \qquad P(\underset{\sim}{Y}_i' A_i \underset{\sim}{Y}_i \le c_i, i = 1, \ldots, k) \ge \prod_{i=1}^{k} P(\underset{\sim}{Y}_i' A_i \underset{\sim}{Y}_i \le c_i)$$

for all positive constants $c_1, \ldots, c_k$.

This is a generalization of Sidak's (1967) inequality in the sense that a symmetric interval generalizes to a symmetric spheroid in higher dimensions and a $n(0,1)$ random variable generalizes to a $n(\underset{\sim}{0}, I)$ random vector. Corollary 2 states that the chi squared random variables $\underset{\sim}{Y}_i' A_i \underset{\sim}{Y}_i$ are positively orthant dependent. Stronger forms of dependency such as associativity, positive regression dependence and monotone likelihood ratio dependence as discussed for example in Dykstra, Hewett, and Thompson (1973) do not hold in general for these chi squared variables.

However, when $k = 2$, more can be said about the nature of the inequality in Corollary 2 as indicated in the following theorem.

<u>THEOREM 2</u>. If $\begin{pmatrix} \underset{\sim}{Y}_1 \\ \underset{\sim}{Y}_2 \end{pmatrix}$ is $n(\underset{\sim}{0}, \Sigma)$, where Σ may be partitioned as

$$\Sigma = \begin{pmatrix} I_{11} & \Sigma_{12} \\ \Sigma_{21} & I_{22} \end{pmatrix},$$

then for all c_1 and c_2

(2.7) $$P(\underset{\sim}{Y}_1'\underset{\sim}{Y}_1 \le c_1,\ \underset{\sim}{Y}_2'\underset{\sim}{Y}_2 \le c_2)$$

is a nondecreasing function of the characteristic roots of $\Sigma_{12}\Sigma_{21}$.

PROOF. There exist orthogonal matrices Q_1 and Q_2 such that

$$Q_1\Sigma_{12}Q_2' = \operatorname{diag}(\rho_1^{1/2}, \ldots, \rho_m^{1/2})$$

where $\rho_1,\ldots,\rho_m$ are the characteristic roots of $\Sigma_{12}\Sigma_{21}$. Since we may express (2.7) as

(2.8) $$\underset{\underset{\sim}{Y}_1^*,\underset{\sim}{Y}_2^*}{E} \; P(Y_{11}^2 \le c_1 - \sum_{i=2}^{p_1} Y_{i1}^2,\ Y_{12}^2 \le c_2 - \sum_{i=2}^{p_2} Y_{i2}^2 \mid \underset{\sim}{Y}_1^*,\underset{\sim}{Y}_2^*)$$

where $(\underset{\sim}{Y}_1^*,\underset{\sim}{Y}_2^*)$ denote the random vectors $(\underset{\sim}{Y}_1,\underset{\sim}{Y}_2)$ with the first components removed, if we increase only ρ_1 , it is well known that the integrand, and hence the whole expression must increase. Since ρ_1 is not special, the theorem easily follows.

In attempting to extend Corollary 2 to ellipsoidal regions rather than spheroids, difficulties are encountered. However, by putting a rather stringent condition on the covariance structure, a result rather similar to Khatri's (1967) result, though stated differently, is possible.

<u>THEOREM 3</u>. Let $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1',\ldots,\underset{\sim}{Y}_k')$ denote a $kp \times 1$ random vector possessing an $n(\underset{\sim}{0},\ B \otimes \Sigma)$ distribution where

$$(2.9)\qquad B \otimes \Sigma = \begin{pmatrix} b_{11}\Sigma & b_{12}\Sigma & \cdots & b_{1k}\Sigma \\ \vdots & b_{22}\Sigma & & \vdots \\ \vdots & & \ddots & \vdots \\ \cdot & \cdots & & b_{kk}\Sigma \end{pmatrix}$$

is the Kronecker product of B and Σ. Then if A is a symmetric p.d. matrix,

$$(2.10)\qquad P(\underset{\sim}{Y}_i' A^{-1} \underset{\sim}{Y}_i \le c_i,\ i=1,\dots,k) \ge \prod_1^k P(\underset{\sim}{Y}_i' A^{-1} \underset{\sim}{Y}_i \le c_i)$$

for all positive constants $c_1, \dots, c_k$.

PROOF. We may assume WLOG that B and Σ are full rank. Then, by using the nonsingular matrix Q such that

$$Q \Sigma Q' = I \quad \text{and} \quad Q A Q' = \operatorname{diag} \Lambda = (\lambda_i)$$

where λ_i are the roots of $|A - \lambda\Sigma| = 0$, we may express (2.10) as

$$(2.11)\qquad P(\underset{\sim}{Y}_i'\Lambda^{-1}\underset{\sim}{Y}_i \le c_i,\ i = 1,\dots,k) \ge \prod_1^k P(\underset{\sim}{Y}_i'\Lambda^{-1}\underset{\sim}{Y}_i \le c_i)$$

where $\underset{\sim}{Y}$ has covariance matrix $B \otimes I$. The augmented matrix

$$B^*\otimes I = \begin{pmatrix} b_{11}I & \cdots & & b_{1k}I & b_{1k}I \\ \vdots & b_{22}I & & & \\ \vdots & & \ddots & b_{kk}I & (1-\varepsilon)b_{kk}I \\ \vdots & & & & \\ \cdot & \cdots & & & (1-\varepsilon)b_{kk}I \end{pmatrix}$$

will be p.d. if $\varepsilon > 0$ is sufficiently close to 0. Thus if $(\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k', \underset{\sim}{Z}')'$ is $n(\underset{\sim}{0}, B^* \otimes I)$, the left side of (2.11) is expressable as

$$\underset{\underset{\sim}{Z}}{E}\ P(\underset{\sim}{Y}_i'\Lambda^{-1}\underset{\sim}{Y}_i \le c_i, i=1,\ldots,k-1|\underset{\sim}{Z})\cdot P(\underset{\sim}{Y}_k'\Lambda^{-1}\underset{\sim}{Y}_k \le c_k|\underset{\sim}{Z}) .$$

However, by the diagonal structure of Λ^{-1} and the conditional covariance structure of $\underset{\sim}{Y}$ given $\underset{\sim}{Z}$, it follows from Lemma 1 that each factor of the integrand is a non-increasing function of Z_1 when $Z_2, \ldots, Z_p$ are held fixed. Repeated applications of Lemma 3 imply the desired result.

One might hope that Theorem 3 would extend in certain situations to the case when Λ is *random. Under the right* conditons, this does indeed happen; the right conditions being that Λ possess a central Wishart distribution with covariance matrix Σ and that Λ be independent of $\underset{\sim}{Y}$.

<u>COROLLARY 3</u>. If $\underset{\sim}{Y}' = (\underset{\sim}{Y}_1', \ldots, \underset{\sim}{Y}_k')$ is $n(\underset{\sim}{0}, B\otimes\Sigma)$, S has a central Wishart distribution with covariance matrix Σ, (S is $W(\nu,\Sigma)$), and $\underset{\sim}{Y}$ and S are independent, then

(2.12) $P(Y_i'S^{-1}Y_i \le c_i,\ i=1,\ldots,k) \ge \prod_1^k P(Y_i'S^{-1}Y_i \le c_i)$ for all positive constants $c_1, c_2, \ldots, c_k$.

<u>PROOF</u>. Clearly we may assume $\Sigma = I$ WLOG. Then the conditional distribution of $(\underset{\sim}{Y}_1'S^{-1}\underset{\sim}{Y}_1, \ldots, \underset{\sim}{Y}_k'S^{-1}\underset{\sim}{Y}_k)$ given

S depends upon S only through its characteristic roots $\underset{\sim}{\psi} = (\psi_1, \ldots, \psi_p)$.

By Theorem 3

$$P(\underset{\sim}{Y}_i' S^{-1} \underset{\sim}{Y}_i \le c_i,\ i = 1, \ldots, k)$$

$$= \underset{\underset{\sim}{\psi}}{E}\ P(\underset{\sim}{Y}_i' S^{-1} \underset{\sim}{Y}_i \le c_i,\ i = 1, \ldots, k \mid \psi)$$

$$\geq \underset{\underset{\sim}{\psi}}{E} \prod_{1}^{k} P(\underset{\sim}{Y}_i' S^{-1} \underset{\sim}{Y}_i \le c_i \mid \psi) \ . \tag{2.13}$$

Since the charactertic roots of a Wishart matrix with identity covariance matrix are stochastically increasing in sequence (Dykstra and Hewett (1978)), and since each factor of the integrand in (2.14) is nondecreasing in ψ_i , Theorem 1 of Dykstra, Hewett and Thompson (1973) preserves the desired inequality when the product sign is brought outside the expectation sign.

3. APPLICATIONS.

Since the right side of the (2.6) is just the product of central chi-squared probabilities, Corollary 2 essentially states that chi-squared random variables which are quadratic forms of a multivariate normal vector are always positively orthant dependent as defined in Dykstra et. al. (1973). However an example of Sidak's (1971) shows that stronger forms of positive dependence like "association", stochastically increasing in sequence" and "positively likelihood ratio dependence" do not hold in general. However, positive orthant dependence does imply non-negative correlations by the expression for the covariance given in Lehmann (1966).

(a) Corollary 2 seems somewhat related to the bivariate chi-squared inequality given by Jensen (1969). However Jensen's inequality, while two-sided, only hold for $k = 2$, equal degress of freedom, and identical intervals. Moreover, since bivariate chi-squared random variables defined in this manner are conditionally independent and identically distributed as shown by Shaked (1977), Jensen's inequality would also hold for any Borel set.

(b) Corollary 2 implies that the product of the marginal c.d.f.'s of multivariate chi-squared random variables serves as a lower bound for the joint c.d.f. of the random variables. A similar statement applies for multivariate F random variables such as discussed by Schuurmann, Krishnaiah and Chattopadhyay (1975). This follow by conditioning on the independent denominator, applying Corollary 2, and then using Lemma 3.

(c) If goodness of fit statistics are defined on different co-ordinates of multivariate data, then the asymptotic chi-squared distributions, assuming the null hypotheses to be true, will satisfy the inequality in Corollary 2. Thus if one rejects the hypothesis that all univariate hypotheses are true whenever a univariate hypothesis is rejected, he will, asymptotically, have a conservative estimate of the significance level if he treats the individual tests as being independent.

(d) An application of Corollary 3 involves simultaneous inference in multivariate linear regression. That is, let Y be a $N \times p$ data matrix of N independent observations on p responses, X be a $N \times q$ design matrix of fixed known independent variables, B be a $q \times p$ matrix of parameters and E be a $N \times p$ matrix of random errors whose rows are distributed independently as $n(\underset{\sim}{0}, \Sigma)$ random vectors. It is well known from least squares theory that the estimator of B which minimizes $\mathrm{Tr}[(Y - XB)'(Y - XB)]$ is given by

$$\underset{q\times p}{\hat{B}} = (X'X)^{-1}X'Y .$$

assuming $X'X$ is of full rank. The distribution of the row-wise rolled out version of $\hat{B}$ is then multivariate normal with covariance matrix $(X'X)^{-1} \otimes \Sigma$ and mean equal to the row-wise rolled out B.

Moreover, $Q_E = Y'Y - \hat{B}'X'X\hat{B}$ is independent of $\hat{B}$ and possesses a central Wishart distribution with $N - q$ degrees of freedom and covariance matrix Σ. Thus if we construct ellipsoidal confidence regions for the i^{th} row of B based on the i^{th} row of $\hat{B}$ and Q_E from the expression

$$P(\hat{\underset{\sim}{B}}_i - \underset{\sim}{B}_i)'Q_E^{-1}(\hat{\underset{\sim}{B}}_i - \underset{\sim}{B}_i) \le c_i) = 1 - \alpha_i ,$$

then Corollary 3 guarantees that the confidence coefficient for all the ellipsoidal regions to contain the respective parameters must be at least $\Pi(1 - \alpha_i)$. This generalizes the known comparable result for univariate linear regression.

(e) Siotani (1959) is concerned with

$$\hat{T}_{MAX} \leq \max_h \{\underset{\sim}{Z}_h' S^{-1} \underset{\sim}{Z}_h\}$$

where $\underset{\sim}{Z}_1, \ldots, \underset{\sim}{Z}_N$ is a random sample from a $n(\underset{\sim}{0}, \Sigma)$ distribution and S possesses an independent $W(\nu, \Sigma)$ distribution. Siotani approximates the distribution of $\hat{T}^2_{MAX}$ by using Bonferroni inequalities. However the product inequality given in Theorem 3 will be closer to the true probabilities and hence could be used to improve Siotani's approximations.

REFERENCES

Anderson, T. W. (1955). The integral of a symmetric unimodal function over a symmetric convex set and some probability inequalities. Proc. Amer. Math. Soc. 6, 170-176.

Anderson, T. W. (1958). An introduction to multivariate statistical analysis. Wiley, New York.

Das Gupta, S., Eaton, M. L., Olkin, I., Pearlman, M., Savage, L. J. and Sobel, M. (1972). Inequalities on the probability content of convex regions for elliptically contoured distributions. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Univ. of Calif. Press, Berkeley, Calif. Vol. II, 241-265.

Dunn, Olive Jean (1958). Estimation of the means of dependent variables. Ann. Math. Statist. 29, 1095-1111.

Dykstra, R. L., Hewett, J. E. and Thompson, W. A. Jr. (1973). Events which are almost independent. Ann. Statist. 1, 674-681.

Dykstra, R. L. and Hewett, J. E. (1978). Positive dependence of the roots of a Wishart matrix. Ann. Statist. 6, 235-238.

Jensen, D. R. (1969). An inequality for a class of bivariate chi-square distributions. Jour. of Amer. Statist. Assoc. 64, 333-336.

Khatri, C. G. (1967). On certain inequalities for normal distributions and their applications to simultaneous confidence bounds. Ann. Math. Statist. 38, 1853-1867.

Khatri, C. G. (1970). Further contributions to some inequalities for normal distributions and their applications to simultaneous confidence bounds. Ann. Inst. Statist. Math. 22, 451-458.

Kimball, A. W. (1951). On dependent tests of significance in an analysis of variance. Ann. Math. Stat. 22, 600-602.

Lehmann, E. L. (1966). Some concepts of dependence. Ann. Math. Statist. 37, 1137-1153.

Schuurmann, F. J., Krishnaiah, P. R. and Chattopadhyay, A. K. (1975). Tables for a multivariate F distribution. Sankya, Series B, 37, 308-331.

Scott, Alastair (1967). A note on conservative confidence regions for the mean of a multivariate normal. Ann. Math. Statist. 38, 278-280.

Shaked, Moshe (1977). A concept of positive dependence for exchangeable random variables. Ann. Statist. 5, 505-515.

Sidak, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. J. Amer. Statist. Assoc. 68, 734-739.

Sidak, Z. (1968). On multivariate normal probabilities of rectangles: their dependence on correlations. Ann. Math. Statist. 39, 1425-1434.

Sidak, Z. (1971). On probabilities of rectangles in multivariate Student distributions: their dependence on correlations. Ann. Math. Statist. 42, 169-175.

Sidak, Z. (1973). A chain of inequalities for some types of multivariate distributions, with nine special cases. Apl. Mat. 18, 110-118.

Sidak. Z. (1975). A note on C. G. Khatri's and A. Scott's papers on multivariate normal distributions. Ann. Inst. Statist. Math. 27, 181-184.

Siotani, M. (1959). The extreme value of the generalized distances of the individual points in the multivariate normal sample. Ann. Inst. Statist. Math. 10, 183-203.

Slepian, D. (1962). The one-sided barrier problem for Gaussian noise. Bell System Tech. J. 41, 463-501.

Tong, Y. L. (1970). Some probability inequalities of multivariate normal and multivariate t . J. Amer. Statist. Assoc. 65, 1243-1247.